

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2001, 15:52:39 ; Search time 170.72 Seconds
(without alignments)
18.693 Million cell updates/sec

Title: US-09-331-631A-7_COPY_34_80

Perfect score: 258
Sequence: 1 YERDPRQOYECCORCESEATERRERQRCERERKQROQEE 47

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	258	100.0	566	2 S22477	vicillin precursor
2	124	48.1	588	1 FMCNAB	alpha-globulin B p
3	124	48.1	605	2 S06398	alpha-globulin typ
4	123	47.7	509	2 S08059	alpha-globulin typ
5	95	36.8	1038	2 T02634	rep protein homolo
6	90	34.9	810	2 T44430	protein PVI100 (Imp
7	87	33.7	429	2 S29565	apolipoprotein A-I
8	86	33.3	47	2 JCS557	arginine/glutamate
9	86	33.3	1898	1 A45973	trichomyalin - hum
10	85	32.9	551	2 S51941	prunin I precursor
11	84.5	32.8	1407	1 S28589	trichomyalin - rad
12	82	31.8	608	2 T06632	hypothetical prote
13	82	31.8	1027	2 T46481	hypothetical prote
14	82	31.8	1233	2 T30989	serine/threonine p
15	81	31.4	613	2 S27770	hypothetical prote
16	80	31.0	1351	2 C71607	hypothetical prote
17	79.5	30.8	401	2 A47141	apolipoprotein A-I
18	79.5	30.8	411	2 T29475	hypothetical prote
19	79	30.6	743	2 A29232	10k melarila anti
20	79	30.6	749	2 T37356	epithelial microtu
21	79	30.6	947	2 T23107	epithelial prote
22	78.5	30.4	1905	2 T18267	multidrug resist
23	78.5	30.4	948	2 T24445	hypothetical prote
24	78	30.2	905	1 R68YSS	regulatory prote
25	78	30.2	1403	2 S24548	homeotic protein p
26	77.5	30.0	502	2 S61935	SKS1 protein - yea
27	77	29.8	81	2 T48398	hypothetical prote
28	77	29.8	339	1 TWHU2D	transcription init
29	77	29.8	646	2 D82493	conserved hypothet

30	77	29.8	772	2	I50463	protein kinase - C
31	77	29.8	858	2	S15762	neurofilament trip
32	76	29.5	292	2	JF0233	tropomyosin-I - scal
33	76	29.5	648	1	JQ1150	protein kinase (EC
34	76	29.5	890	1	F1EC2	translation initia
35	76	29.5	1457	2	T14577	protein kinase yak
36	76	29.5	2019	2	T27702	hypothetical prote
37	75.5	29.3	1023	2	S12519	glutactin - fruit
38	75.5	29.3	1737	2	A59235	unconventional myo
39	75	29.1	241	2	S44893	ZK1236.7 protein -
40	75	29.1	411	2	S29129	calreticulin precu
41	75	29.1	758	2	S54522	hypothetical prote
42	75	29.1	966	2	S25365	Cyc6 protein - yea
43	75	29.1	1038	2	S52522	hypothetical prote
44	75	29.1	1069	2	T00377	KIAA0642 protein -
45	74.5	28.9	1549	1	A40691	trichomyalin - she

ALIGNMENTS

```

RESULT 1
S22477
vicillin precursor - cacao
C:Species: Theobroma cacao (cacao)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 17-Mar-1999
C:Accession: S22477; S22478; S18105; S22050
R:McHenry, L.; Fritz, P.J.
Plant Mol. Biol. 18, 1173-1176, 1992
A:Title: Comparison of the structure and nucleotide sequence of vicillin genes of coco
A:Reference number: S22477; M01D:92288309
A:Accession: S22477
A:Molecule type: DNA
A:Residues: 1-566 <MC6>
A:Cross-references: EMBL:X62625
A:Accession: S22478
A:Molecule type: mRNA
A:Residues: 1-452 <MC2>
A:Cross-references: EMBL:X62626
C:Genetics:
A:introns: 211/1; 269/3; 296/3; 391/3; 502/1
C:Superfamily: glycinn
C:Keywords: seed; storage protein
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-566/Product: vicillin #status predicted <MAT>

Query Match          100.0%  Score 258;  DB 2;  Length 566;
Best Local Similarity 100.0%  Pred. No. 4e-17;
Matches 47;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

OY 1 YERDPRQOYECCORCESEATERRERQRCERERKQROQEE 47
DB 34 YERDPRQOYECCORCESEATERRERQRCERERKQROQEE 80

RESULT 2
FMCNAB
alpha-globulin B precursor (clone C72) - upland cotton
A:Alternate names: seed storage protein; vicillin precursor
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: A30838; S06911
R:Chian, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.
Plant Mol. Biol. 7, 475-489, 1986
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII
A:Reference number: A30838
A:Accession: A30838
A:Molecule type: mRNA
A:Residues: 1-588 <CHL>
A:Cross-references: GB:M16891; NID:g167374; PIDN:AA33071.1; PID:g167375
R:Chian, C.A.; Borrolo, R.; Kamalay, J.A.; Dure III, L.

```

Plant Mol. Biol. 9, 533-546, 1987
A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Seed
A:Reference number: S06398
A:Accession: S06911

A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-81 <CH2>
C:Comment: This is a seed storage protein.

C:Keywords: glycoprotein; seed; storage protein
F:1-25/Domain: signal sequence #status predicted <StG>
F:26-588/Product: alpha-globulin storage protein #status predicted <Mat>
E:417/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	48.1%	Score 124;	DB 1;	Length 588;
Best Local Similarity	46.3%	Pred. No. 0.0001;		
Matches 19;	Conservative 12;	Mismatches 10;	Indels 0;	Gaps 0

QY 4 DPROQYDQCRCRCESEATEEREDEDQCEBCRCERYKKEQORQ 44
 ||::||:||||: ::||:||||: |: :||: :||
 Db 36 DPKRYEDCRRRCRCEWDFRGQKEQDQCESSCKSQYGEKQDQ 76

RESULT 3
S06398
alpha-globulin type A precursor - upland cotton

C:Species: *Gossypium hirsutum* (upland cotton)
C:Date: 31-Mar-1990 #sequence-revision 31-Mar-1990 #text-change 30-Sep-1993
C:Accession: S06398
R:Chan, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.

A;Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Se
A;Reference number: 506398
A;Accession: 506398
A;Status: not compared with conceptual translation

A:Residues: 1-605 <CHL>
C:Superfamily: glycinin
E:1-24/Domain: signal sequence #status predicted <Sto>
F:25-605/Product: alpha-globulin type A #status predicted <Mat>

Query Match	48.1%;	Score 124;	DB 2;	Length 605;
Best Local Similarity	43.5%;	Pred. No. 0.00011;		
Matches	20;	Conservative 13;	Mismatches 13;	Indels 0;
			Gaps	0

```
QY      2 EDDPROQVEQCQRCESEATEEREQECCQRCEREXKEQQRQEEE 47
          | :|::||| :|::| :| :|::||| :| :|::| :| :|::| :| :|:
Db      34 EDDPQRYEDCRKRCQLTREGQTEQDKCEDRSLETQLKEEQORDGED 79
```

RESULT	4
508059	

alpha globulin type 2 precursor (unbound) upland cotton (irrigation)
 N:Alternate names: seed storage protein
 C:Species: *Gossypium hirsutum* (upland cotton)
 C>Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 30-Sep-1993
 #Accession: S08059

A:Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. Seed
A:Reference number: 506398
A:Accession: 508059
Nikhar, V.R.; Dholak, A.P.; Kadamaty, C.R.; Dale L.L., L.
Plant Mol. Biol. 9, 533-546, 1987

A: Molecule type: DNA
A: Residues: 1-509 <CHL>
C: Superfamily: glycinin

Query Match	47.78;	Score 123;	DB 2;	Length 509;
Best Local Similarity	40.98;	Pred. No. 0.00011;		

	Matches	18;	Conservative	16;	Mismatches	8;	Indels	2;	Gaps	1
QY	4	DPRQYEQQRCESEATEEREDQECQRCSEFYEQDQROEEF	47							
		:: : : : :: : :::: : :								
Pb	1	DPRRYEQQDECRQQ--EERRQPCCQQRCLKRFDEQDSORQ	42							

RESULT	5
T02634	

C:Species: Dictyostellium discoideum
C:Date: 24-Mar-1999 #sequence-revision 24-Mar-1999 #text-change 29-Oct-1999
C:Accession: T02634
R:Rieben Jr., W.K.; Gonzales, C.M.; Gonzales, S.T.; Pilkington, K.J.; Kiyose

A:Title: Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to the Dd
A:Reference number: Z14684; MUID:98198836
A:Accession: T02634
A:Status: preliminary; translated from GE/EMBL/DBJ

A:Residues: 1-1038 <RIE>
A:Cross-references: EMBL:AF000580; NID:g3068582; PIDN:AAC14390.1; PID:g3068583
C:Experimental source: strain WS2162
C:Genetics:

A;Genome: plasmid
A;Mobile element: plasmid ddp5

Query Match	36.8%	Score 95;	DB 2;	Length 1038;
Best Local Similarity	37.0%;	Pred. No. 0.082;		
Matches 17;	Conservative 18;	Mismatches 11;	Indels 0;	Gaps 0

[illegible]

RESULT 6
T44430
protein PV100 [imported] - winter squash

Cidbate: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-20000
CjAccession: T44430
R:Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
C: Biol. Chem. 274, 2563-2570, 1999

A:Reference number: 222767; MUID:9107919
A:Accession: T44430
A:Status: preliminary; translated from GE/EMBL/DBJ
A:Molecule type: mRNA

A:Cross-references: EMBL:AB019195; NID:g3808061; PIDN:BAA34056.1; PID:g3808062
 A:Residues: 1-101 (1AMP)

5 PROOYRQOCORCSFATRREREROCRCRRCRFRFKKFPDORRORFFR 47
Query match 34.9%; Score 90; DB 2; Length 810;
Best local similarity 34.9%; Pred. No. 0.19;
Matches 15; Conservative 11; Mismatches 17; Indels 0; Gaps 0

```

Db      74  PRAEYECRLRCQVAERGVREQRKCEQVCCEERLRERQGRGED 116
          *      *      *      *      *      *      *      *
RESULT  7

```

529565
apolipoprotein A-IV - crab-eating macaque
C1Species: Macaca fascicularis (crab-eating macaque)
C1Date: 06-Jan-1995 #sequence_revison 06-Jan-1995 #text_change 13-Aug-1999
#accession 529565

A-7>Title: Nucleotide sequences of the Maccia fascicularis apolipoprotein C-III and A-I-I

